

In the museum of the Louvre, in Paris, sits “The Raft of the Medusa”, one of the most celebrated paintings of the nineteenth century, the work of Theodore Gericault. Here it is, in reproduction (Fig. 18.1).

Obviously the scene on board this raft is catastrophic. Bodies writhe in anguish, and a piece of cloth is raised in a desperate effort to get help. If you look very closely, you can see that there is a ship far in the distance. Will it see the raft and come to its rescue?

The story behind this painting is worth retelling. In June of 1816, a flotilla of four ships departed France carrying soldiers to Africa. One of them, the *Medusa*, struck a reef in the Atlantic Ocean and suffered serious damage. For 3 days its crew struggled to free the ship, but eventually the *Medusa* began to break apart and sink. The leaders of the expedition—the officers of the ship, the military commanders from France and a high-level government official—took control of the lifeboats, with the intention of saving only themselves. In response, the common soldiers and crew members lashed together a raft from the timbers of the failing ship and convinced their superiors to tie this raft to the lifeboats. It soon became clear that the boats could not make much progress if they had to tow the raft, so the rope connection was cut by one of the officers. The boats rowed to safety while the raft drifted helplessly in the middle of the ocean.

There were 154 people aboard the raft. They had little food or water and hardly any space to move. They were adrift for 11 days before finally being seen by a passing ship. By that time all but 13 had died horribly, including the cloth-waver depicted by Gericault. When the news traveled back to France, it was regarded as a scandal: how could some of the most privileged men in the country turn their backs on those whose work made their riches and power possible in the first place? It led to an upheaval in the government, and the appearance of Gericault’s painting 3 years later was electrifying in an age before photography or rapid communication.

For many at this time, the terrible fate of these crewmen served as a metaphor for the development of the modern economy. With the explosion of new technologies and the rise of new business empires, some were becoming very rich, but most of



Fig. 18.1 The Raft of the Medusa (Gericault) (Source: Wikimedia Commons, available at http://upload.wikimedia.org/wikipedia/commons/1/12/Raft_of_the_Medusa.jpg)

the others, it seemed, were being left behind. Progress for the few was based on cutting the rope.

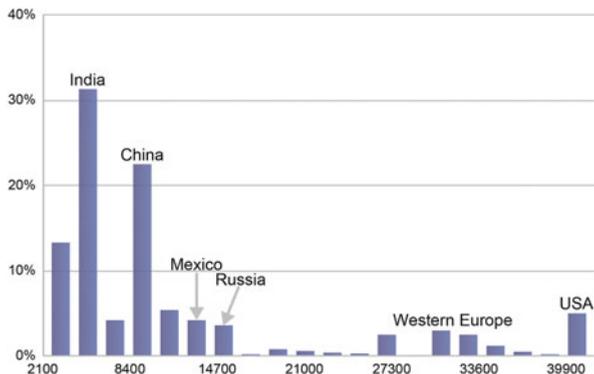
Modern historical research has shown that, in most of western Europe, improvement in living standards for the majority of the population lagged behind economic growth during much of this period. In England, for example, whose propulsive economy was already apparent by the middle of the eighteenth century, it was not until about the middle of the nineteenth that the benefits began to “trickle down” to most of the working class. One reason that so many farmers and artisans left Europe for the New World in this period is that their lives were not getting better at home.

But historians have also shown that, for the most part, the lives of the common people were not getting *worse*. (There are obvious exceptions, of course, such as the famine that struck Ireland during the 1840s.) In this sense, the Medusa parable is misleading. Most workers and farmers in Europe were not aboard a death raft but simply living as they had for generations. Yet it seemed far worse than this because some were doing so much better. In other words, it was not absolute deprivation that had exploded but *inequality*.

Societies have always been unequal to some extent, and certainly those built on slavery or aristocracy are the most unequal of all. Still, the emergence of capitalism has put the issue of inequality in a new light, since wealth is now being created at an unprecedented rate, but not for all. It is the promise of general abundance instead of general scarcity that makes inequality such an important topic today. For the first time, we can imagine that all human beings could live fairly comfortably, even though they currently don't.

This concern shows up in two ways. First, there is an interest in what we might call **general inequality**, the differences between income, wealth and living

Fig. 18.2 Percentages of the global population by gross domestic product per capita, 2010 (Constant 2005 PPP). (Source: World Bank World Development Indicators)



standards across the population. Here the question is, how concentrated is the distribution of these things? Are they enjoyed primarily by just a few, or are they spread out more evenly among everyone? Second, we have become interested in comparisons between certain groups: racial and ethnic groups, men and women and higher and lower castes, for instance. Are some groups more advantaged than others? If so why, and what can be done about it?

Let's begin by surveying economic inequality in the world we live in.

18.1 Some Initial Evidence

It will come as a surprise to no one that this is a highly unequal world, with people in some countries having much higher incomes than others. Figure 18.2 shows the average income per capita for different fractions of the global population, based on the country they live in.

A few clarifications are in order. GDP, gross domestic product, is a measure of the value of everything a country produces during a year, which is also a measure of the income people receive from this production. GDP per capita is therefore approximately the average income in a country, which is not necessarily similar to the median, or typical, income. (A few very high-earners can pull the average up.) PPP is "purchasing power parity", a basis for converting income in one currency to a world standard; we will discuss its measurement in the volume on macroeconomics. Some countries did not report GDP data for 2010; the chart reflects only those that did. (They account for 97.7 % of the world's population.) Finally, the chart does not identify most of the countries, so the ones it does mention "stand in" for the others. For instance, neither India nor China has more than 20 % of the world's population alone; they are joined in their bars by other, smaller countries with approximately the same GDP per capita.

We can see that a large majority of the planet's people live in poorer countries, with average incomes well below \$15,000 (PPP). Surprisingly few live in what

might be termed the world's middle class of countries, with averages in the \$15,000–\$25,000 range. This is a Medusa-like picture.

Seeing the world's income in terms of country averages is incomplete, however. Within every country there are large differences among citizens; some fortunate residents of the poor countries are quite rich, while some in the rich countries are relatively poor. What would we find if we ignored national differences and looked at the world as a single population, with each individual holding his or her own place in line? Branko Milanovic, a World Bank economist, has spent many years trying to calculate this, and he finds that world inequality in this sense is equal to or greater than inequality in any single nation that has ever been measured. The world Gini coefficient, which we will define later in this chapter, is approximately 0.650, meaning that the world is roughly two-thirds of the way from perfect equality to the perfectly unequal situation in which one person would get 100 % of all income. To illustrate this, he writes:

the top 5 percent of individuals in the world receive about 1/3 of total world (PPP-valued) income, and the top 10 percent one-half. If we take the bottom 5 and 10 percent, they receive respectively 0.2 and 0.7 percent of world total income. This means that the ratio between the average income received by the richest 5 percent and the poorest 5 percent of people in the world is 165 to 1. (Milanovic, 2006, p. 16)

If everyone worked an average of 2,000 hours per year (50 weeks times 40 hours per week), the richest 5 % would make more in about a day and a half (12 h) than the poorest 5 % would make all year.

As much as we might think of ourselves as citizens of the world, however, most of us spend most of our time among others of our own nationality, and this is the source for the usual comparisons we make when we think about how well-off we are. Also, the political institutions under which we live cannot currently control global inequality, but they have many policies that make income within our countries more or less equally distributed. So what is the level of income inequality at the national level?

Table 18.1 summarizes this for a sampling of countries, poor, rich and in-between. The measure is once again the temporarily mysterious Gini coefficient. For now, the only important feature of this statistic is that a higher value signifies greater inequality; it ranges between 0 and 1 (Table 18.1).

When reading this table, bear in mind that more equality alone does not mean “better”; there are also large differences in average income to consider. Income in Indonesia is distributed more equally than in Britain, but most residents of Britain have much higher incomes than most Indonesians. If you want to make comparisons, the logical ones would be between countries at approximately the same levels of average income—Indonesia and Cote d'Ivoire, for example, or the United Kingdom and Sweden. Certain generalizations seem to hold:

- Poorer countries tend to have more inequality than richer ones, but there are also many exceptions.
- Inequality is high in Latin America and the Caribbean.

Table 18.1 Gini coefficients for selected countries

Country	Year	Gini coefficient
Algeria	1995	0.353
Indonesia	2002	0.343
Argentina	2003	0.528
Iran	1998	0.441
Bangladesh	2000	0.318
Ireland	200	0.343
Bolivia	2002	0.601
Israel	2001	0.392
Brazil	2003	0.58
Kenya	1997	0.449
Canada	2000	0.326
Korea (South)	1998	0.16
Chile	2000	0.576
Mexico	2002	0.495
China	2001	0.447
Nigeria	2003	0.436
Cote d' Ivoire	2002	0.446
Russia	2000	0.456
Denmark	1997	0.247
South Africa	2000	0.578
Egypt	2000	0.344
Sweden	2000	0.25
France	1995	0.327
Tanzania	2001	0.346
Germany	2000	0.283
Thailand	2002	0.42
Guatemala	2002	0.551
Turkey	2003	0.436
Haiti	2001	0.592
United Kingdom	1999	0.36
Hungary	2002	0.269
United States	2000	0.408

Source: World Bank, World Development Indicators

- English-speaking countries tend to have relatively high levels of inequality compared to others at their income levels, and the United States has the highest of any of these.

In any case, it is clear that inequality varies enormously across countries, measured, as in this case by Gini coefficients, by a factor of more than two to one. The level of inequality in the United States is more than half again as high as in Sweden; the level in Brazil more than twice.

These simple national comparisons leave out a lot of detail. They don't tell us which social groups tend to benefit the most, or where in the distribution (the top or

Fig. 18.3 Income inequality among US families, 1947–2010, measured by Gini coefficients. (Source: Federal Reserve Bank of St. Louis)



the bottom) inequalities tend to be the strongest, or what changes are taking place over time. To get a closer reading we have to focus on just one country at a time. Here we will look at the United States, whose government and private research institutions collect excellent economic and social data.

The long-term trend in the US has been toward greater inequality. Figure 18.3 shows changes in the Gini coefficient for family income, which is somewhat greater than individual income, due to the tendency for higher income-earners to live in the same families.

There was a long post-WWII downward trend that reversed itself in the 1970s; since that time inequality has risen in most years. This pattern, particularly the accelerated movement toward inequality since around 1980, can be found in many other countries as well.

Even to those who are familiar with them, Gini coefficients can be a bit opaque, so on the following page are the same data grouped into quintiles (fifths), with the top 5 % broken out of the top quintile (Table 18.2).

The bottom 80 % of the family income distribution all had reduced shares of the total income between 1973 and 2000; the difference went to the top 20 %. Most of the gains of the top quintile were concentrated, in fact, in the top quarter of this group, those who made up the highest 5 % of all families. There has been some stabilization since 2000, although the share of the bottom 40 % continues to fall, and the share received by the top 1 % (not depicted) has risen substantially.

18.2 What About Mobility?

Inequality at any point in time would be more acceptable if there were few barriers in the way of moving up or down the economic ladder. If you were poor and were reading about someone who was fabulously rich, you might think, “I know there’s a big gap between us, but maybe tomorrow I will be the one who’s on top, and others will be looking up at me.” The patron saint of this way of looking at temporary

Table 18.2 Percent of US family income received by quintiles and the top 5 %, 1947–2010

Year	Lowest fifth	Second fifth	Middle fifth	Fouth fifth	Top fifth	Top 5 %
1947	5	11.9	17	23.1	43	17.5
1973	5.5	11.9	17.5	24	41.1	15.5
1979	5.4	11.6	17.5	24.1	41.4	15.3
1989	4.6	10.6	16.5	23.7	44.6	17.9
2000	4.3	9.8	15.4	22.7	47.7	21.1
2007	4.1	9.7	15.6	23.3	47.3	20.1
2010	3.8	9.5	15.4	23.5	47.8	20

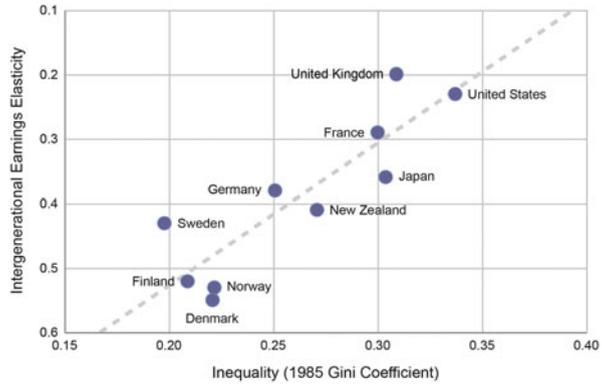
Source: Economic Policy Institute

inequality was Horatio Alger, a popular nineteenth century author in the US, who wrote stirring novels like *Struggling Upward*, *Sink or Swim* and *Bound to Rise*. In each of his creations, Alger presented a young, poor-but-plucky boy (always a boy) who seized opportunity and rose to the ranks of the elite. No doubt his story came true for many young people then as it does today. Of course, there are also those who try with all their might and are unable to overcome the obstacles in their path. The question is not whether there is mobility from bottom to top (and vice versa) in society, but how much.

A modern and more factual depiction of income mobility is given by Fig. 18.4 on the next page. It has been dubbed the “Great Gatsby Curve” in honor of the fictional hero (sort of) of another American novelist, F. Scott Fitzgerald. (*The Great Gatsby* is a penetrating look at the economic elite at the time of its writing in 1925.) As it makes clear, mobility across generations, rather than compensating for an increase in inequality within any single generation, has a tendency to make matters worse.

On the horizontal axis we see our familiar friend, the Gini Coefficient; remember that a higher Gini signifies greater inequality across individuals at a moment in time. On the vertical axis is intergenerational earnings elasticity. What this measures is the percentage change in your expected income today given a percentage change in (in this case) your father’s income a generation ago. For instance, if this elasticity is .3, as it is, approximately, for Germany and New Zealand, if your father’s income was 10 % higher than mine was 20 years ago, on average your income will be 3 % higher than mine today. Not surprisingly, every country in this sample has a positive elasticity of this sort: the advantages of one generation are always handed off, more or less, to the next. But the amount differs enormously. For the countries with the most intergenerational mobility—Denmark, Norway and Finland, 10 % more for the father translates, again on average, to only 2 % more for the child. In the US and the UK, on the other hand, 10 % more in the first generation implies 5 % more in the second: there is still some mobility but far less. What is especially interesting about this curve is that it is a *curve*. The straight line shows the overall trend, which is for countries that are more equal at a given year (in this case 1985) to also have more mobility across generations. The trend has a strong upward slope, and the actual points are not too far from the trend.

Fig. 18.4 Intragenerational inequality and intergenerational mobility. (The Great Gatsby Curve) (Source: Krueger (2012))



(The Gatsby relationship embodied in the trend line explains, on its own, three-fourths of the mobility differences between countries.)

As we think about inequality at one moment in time and across generations, the logic should not be surprising. The wider the difference between income groups, the harder it will be to make the trek from one group to the next. In general, the sort of policies that promote more inequality among the current population will also tend to promote more mobility up and down the ladder for the next generation.

18.3 Inequality by Gender and Race

One of the great revolutions of the twentieth century (it's too soon to say much about the twenty-first) has been the push for greater equality—economic, social and political—between men and women. Figure 18.5 shows the progress, and limits, of this revolution in the ratio of women's to men's hourly earnings. (In making sense of this chart, note that many workers have more than a high school degree but less than a college BA.)

Racial and other barriers are also slowly coming down. You can see how much progress remains to be made by examining the wage ratios by race and ethnicity shown in Fig. 18.6.

Blacks and Hispanics, as groups, fare worse than whites, and the trend over the past three decades is down somewhat. The trend is worse for Hispanics, but it is important to bear in mind that their population has changed dramatically between 1979 and 2007; not only have their numbers increased, but many more are likely to be first-generation residents. Incidentally, when economists analyze data like these, they prefer to adjust the results to control for factors like different levels of education and work experience. For instance, Blacks in the US have somewhat less education on average than Whites. Since, as we have seen, education is linked to earnings, it makes sense to ask how much of the Black-White wage difference appears to come from pure discrimination in the labor market rather than differences in years of education. On the other hand, this approach can lead us to

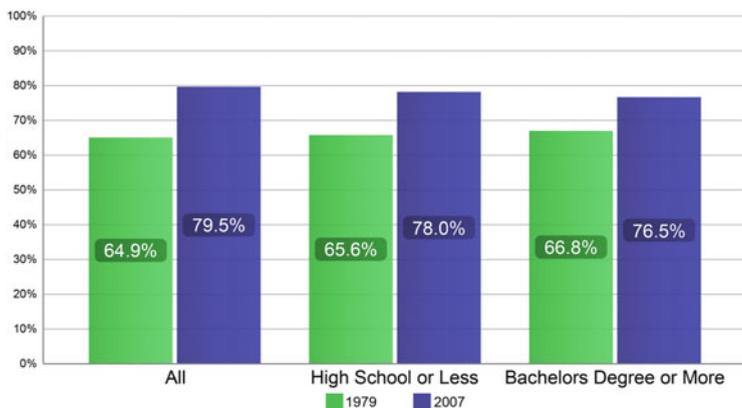


Fig. 18.5 Percent of women’s to men’s average hourly wage, US, 1979 and 2007. (Source: Holzer and Hlavac (2012))

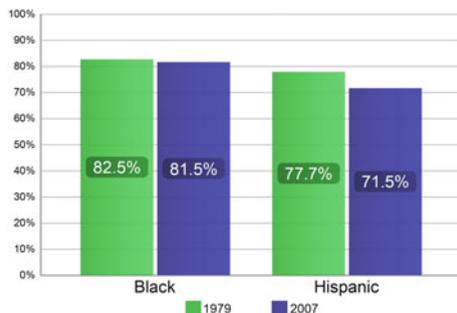


Fig. 18.6 Percent of median US white male hourly wage by race, ethnicity and gender, 2005. (Source: Holzer and Hlavac (2012))

underestimate the overall effect of discrimination on economic inequality. There is plenty of evidence of subtle and not-so-subtle prejudice in the schools, for instance, as well as differences in the quality of schools between inner cities and suburbs. Similarly, differences in overall work experience and the number of years at the current job between men and women are largely the result of different expectations women and men face concerning housework, childcare and taking care of the sick and elderly—themselves manifestations of a larger social inequality.

With these qualifications, and keeping this whirlwind tour of global and US income inequality in mind, let’s take a look at some of the analytical tools economists and others have developed to examine and explain the numbers.

18.4 Measuring Inequality

The simplest approaches to quantifying the extent of income inequality in a population are to show the amount of income received by a particular portion of the population or by the person whose income places him or her at a particular point within the distribution. We have already seen examples of the first of these: the percent of income according to quintile (fifth), top 5 % and top 1 %. Very similar in spirit is to compare two people at different percentage point levels, for instance at the 90 % percentile (where exactly 90 % of the population lies below) and the 50 % percentile (the median). This is very easy to visualize because these are real incomes received by real, flesh-and-blood people.

The problem with comparisons of individuals and groups is that they capture only a part of the degree of inequality. For instance, in Table 18.2 we see that in 2005 the top fifth of US families received 48.1 % of the total income, but how equally distributed was the income within this fifth? We have one breakout, the upper 5 %, but this tells us nothing about the upper 1 %, 2 %, etc. Similar criticisms could be made about the data for the other fifths: they are quite broad and cancel out too much detail. In the end, the overall level of inequality is determined by the distribution at *every* point along the way from top to bottom. How can we quantify this?

There are several methods, but the most commonly encountered is the **Gini coefficient**. Before we can show how it is calculated, however, we have to first explore a related concept, the **Lorenz Curve**. Suppose we come upon a society in which income is distributed in a perfectly equal way, so that each individual (or family or some other unit of measurement) gets the same as every other. We could picture this in a chart like Fig. 18.7a. The horizontal axis measures the cumulative percentage of the population. To make things simple, consider two such percentages, 33 and 67 %. Since each has an income perfectly corresponding to its population size, 33 % of the people have 33 % of the income and 67 % have 67 %. This is shown on the vertical axis, which measures the cumulative percentage of income. In this way, by tracing combinations of population and income percentages we could draw a line from the lower left corner to the upper right. It emerges from the horizontal axis at a 45° angle and ends where 100 % of the population has 100 % of the income.

The other extreme is represented by Fig. 18.7b. In this case only one person has all the income. This means that the first third have nothing, the second third nothing, and cumulative income remains zero until the final individual is counted. The line running along the horizontal axis and then leaping up the vertical axis at the far end encompasses all the population/income points.

Both purple lines are examples of **Lorenz Curves**, but neither is very likely. A more typical case is given by Fig. 18.8, which is taken from the 2005 data for quintiles in Table 18.2.

We begin at 0 where there are no people and no income. Our first stopping point is at (20, 4.0), since the bottom 20 % of the population hold 4.0 % of the income. The next 20 % add an additional 9.6 %, so the *cumulative* income held by the

Fig. 18.7 Two extreme Lorenz Curves. (a) In a perfectly equal distribution, each percentage of the population (such as 33 or 67 %) receives exactly the same percentage of income. If the percent of population is measured along the horizontal axis and the percent of income along the vertical axis, a 45° angle line (in purple) includes all the population/income points. (b) In a perfectly unequal distribution, the percentage of income remains zero until we include the last individual, and then it goes directly to 100. Thus the line in purple includes all the population/income points

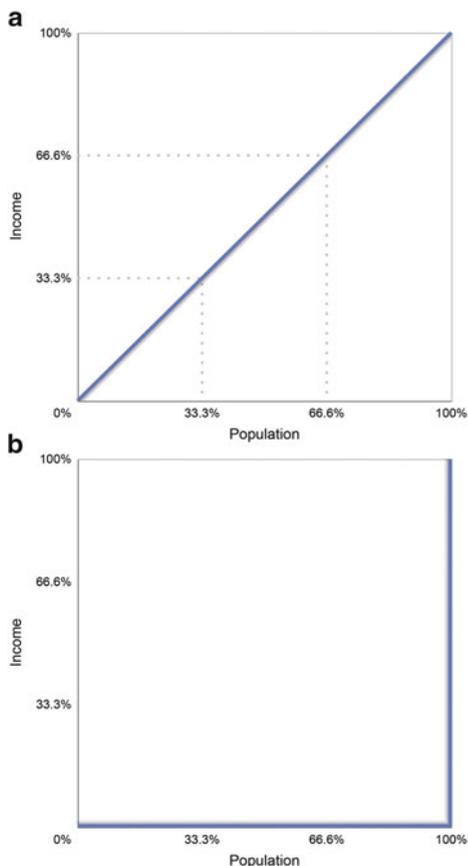


Fig. 18.8 A Lorenz Curve for US family income distribution, 2005. The purple curve is drawn from four points representing the cumulative income percentages received by the bottom 20, 40, 60 and 80 % groups, plus the zero-zero and 100–100 points, which always hold

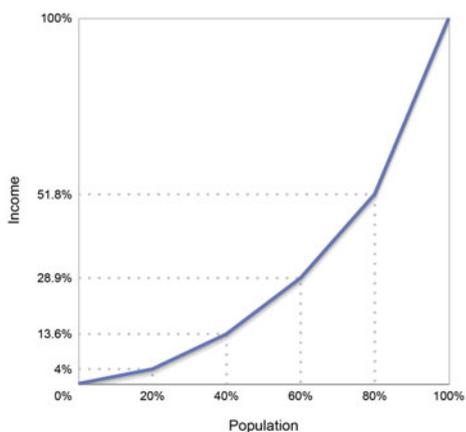
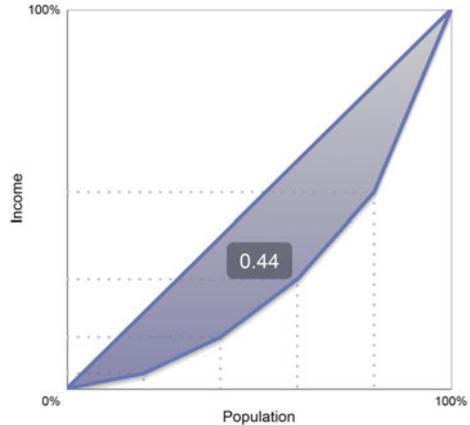


Fig. 18.9 Calculating the Gini coefficient for US family income distribution, 2005. The Gini coefficient expresses the area between the 45° line (perfect equality) and the Lorenz Curve (actual equality) as a fraction of the total area under the 45° line. This shaded area is a bit less than half the potential triangle



bottom 40 % is 13.6 %. This method continues for the next two quintiles, and we finally arrive at 100 % of the population receiving 100 % of the income. The reason the line looks a bit jagged is that it was drawn using only the five quintiles. Thanks to the size and accuracy of annual (and even monthly) census data, however, we could, without too much difficulty, fill in as many intermediate points as we might like, making the line smooth. In this way, our Lorenz curve would reflect the distribution of income across every household.

Now we are ready to consider a technique that reduces the Lorenz curve to a single number. How could we compare the distribution in Fig. 18.8 to a perfectly equal distribution, such as in Fig. 18.7a? Figure 18.9 shows both of them in one diagram, with the area between them shaded in. The more unequal the true distribution, the larger this area, as the Lorenz curve begins to look more like the extreme version in Fig. 18.7b. The Gini coefficient is simply the ratio of this shaded area to the entire area below the 45° line. Perfect equality would be zero (no shaded area), perfect inequality 1. In the US case, as we saw in Fig. 18.3, the Gini is about 0.44. This means that if we draw the Lorenz curve very accurately and base it on each potential data point (and not just the quintiles), the shaded area would be a little less than half the area of the entire triangle making up the bottom half of the box. This is why we can say that inequality in the US lies about halfway between complete equality and complete inequality.

What information do we lose when we reduce an entire diagram to a single number? The relative extent of inequality along different portions of the whole distribution. For instance, the Gini coefficient for the US does not tell us how much of the inequality is due to differences among low- and middle-income families, and how much because of inequality at the very top. In the US case, the top households have a big effect on the entire measurement; for instance, the top one-hundredth of one percent of individuals received about 8 % of all income in 2005. This would be reflected in extreme steepness of the Lorenz Curve as it approaches the upper right-hand corner.

Now that you have a better sense of how the Gini coefficient is calculated and what it means, look again at Table 18.1 and Fig. 18.3.

18.5 The Functional Distribution of Income

Thus far we have thought about the issue of inequality solely in terms of populations, but economists tend to approach the question in a rather different way. Beginning with the classical political economists of the eighteenth century, like Adam Smith, there has been a tendency to think about the sources of income, stemming from the roles different people play in the economy. For Smith and his contemporaries, the big question was, what determines the distribution of society's income into wages (to workers), profits (to owners of capital) and rents (to owners of land and natural resources). To this modern economists would add at least one other, returns to scarce human capital (highly valued skills and abilities), and sociologists would consider the reward to status, particularly to achieving a position high in corporate or similar hierarchies.

Profit has already made an appearance several times in this book. It first came to us in its average form, as the opportunity cost of capital—the return an investor could reasonably expect to get on a typical financial investment. Then we saw the role it plays in the theory of the firm, as the main goal of management (the sole goal in shareholder-driven financial systems). And it was seen to rise or fall based on the bargaining power of labor. We could say, then, that we already know quite a bit about profit, its origin and significance. Nevertheless, from the standpoint of income distribution there is something more to say.

The effect of profits on income distribution derive from two factors, the extent to which ownership of capital is concentrated in society, and the return on capital compared to the return on labor and other factors of production.

In all societies capital is very unequally owned. Once again, we can let the United States, with its rich sources of statistical data, stand in for other countries. If we take the broadest definition of wealth, the value of all household assets (stocks, bonds, pension accumulations, housing) minus all debts, we find that in 2004 the top 20 % of households held about 85 % of net worth, and the top 1 % just over a third. If we look only at net financial assets (stocks and bonds), the holdings of the top 1 % rise to over 42 %.

Returns on capital fluctuate greatly from year to year based on the performance of the financial markets. In good years, like the late 1990s in the US, profits shoot up. In bad years, like the early 2000s and the dismal year 2008, they are depressed. 2006, for instance, was between these two extremes. In that year a little more than half of all business and profit income went to the top 1 % of households, who received just over half their income from these sources. Overall, business and profit income together accounted for about 18 % of all income in society; about 65 % was a return to labor. (The remainder came from a variety of sources, including government transfers.) What this shows is that returns to investment (in one's own business and in financial assets generally) played a modest role in determining

the overall level of income inequality in the US, but a very large role in determining the share taken at the very top.

Because of their interest in the role of human capital, economists have paid close attention to the tendency in the US and other countries toward a wider gap between the earnings of more- and less-educated workers. To take two groups in particular, in 1973, when there were almost four times as many workers whose education had stopped at high school compared to college graduates, the ratio of the average high school hourly wage to those of college grads was about 69 %; in 2005, when there were only half again as many high school-only workers as college grads, it had fallen to 57 %. Similar differences have expanded between other education groups, and not only in the US but in most other industrialized countries.

Organizational standing is most visible in the pay received by CEO's (chief executive officers) of corporations compared to the earnings of those who work under them. The average of this multiple in the US corporate sector has grown from 27 in 1973 to 262 in 2006, although this trend is not nearly as strong in most European countries and Japan. Similar, if not quite as dramatic, pay gaps have also widened further down the hierarchy, for example in the relationship between the earnings of managers and those they supervise.

Although the strengths of their effects vary, all the functional factors point in the same direction, toward a tendency to greater inequality in income across the population. This is one of the most important developments in the world economy during the past few decades, and it is even more striking when set against the long period of *declining* inequality that characterized the 20 years following the end of World War II. Why this has occurred and what, if anything, should be done about it are the questions we turn to next.

18.6 Income Inequality: Explanations and Policy Responses

Very broadly, four different theories have been put forward to explain the sort of evidence we have surveyed in this chapter: (1) "skill-biased technical change", (2) globalization, (3) deregulation and (4) "winner-take-all". Before we go into each one separately, it is important to bear in mind that they are not mutually exclusive; all may capture some part of the truth, and there may be other causes analysts have thus far failed to consider.

1. **Skill-biased technical change.** For anyone who has lived through the technological revolutions of the last 25 years or so, nothing has been as striking as the transformation of the digital computer from a costly behemoth sitting in the basement of large office buildings to the personal tool perched on the desktop, laptop or even in the shirt pocket. The proliferation of computers, along with email and the Internet, has changed how business is done and who does it. It would be surprising if it had no effect on the distribution of income among those with different computer skills or access to these technologies.

Many economists think that computerization is "skill-biased"; it increases the returns to those with greater education or other forms of human capital and

decreases the opportunities for those who are less able to make use of them. Many studies have been done of the relationship between whether a computer is used at work and how much a worker earns; the results are suggestive but not yet conclusive. It is very likely, however, that the computer, by automating many routine tasks, both manual and mental, has changed the labor market in a way that is highly disadvantageous for those with less education.

2. Globalization. By reducing barriers to trade, particularly in manufacturing, but now also in many clerical and service occupations, globalization has intensified competition between the workers of different countries. Many in higher-income countries are seeing their wages decline as they and the companies they work for have to hold their own against increasingly productive workers in lower-wage regions.

At the same time, other workers have found their opportunities increased by globalization. Those with scarce skills are the beneficiaries of heightened competition, a phenomenon that can be seen in fields like finance and entertainment. As the world becomes more tightly integrated, the “world players” can operate on a larger stage. Also, the rules under which globalization is taking place have tended to protect some workers more than others. The result is a growing gap between globalization’s winners and losers.

3. Deregulation. Behind this theory is the presumption that “raw”, unregulated capitalism tends toward very high levels of inequality. If competition is the only force at work, some will succeed fabulously, others will fail dismally, and differences between the better- and worse-off will compound over time. This presumption may not be true, however, or it may be true only some of the time and for some portions of the population, but for now let us assume that it is mostly correct. What would this tell us about current trends toward inequality?

Throughout much of the twentieth century, in most regions of the world, there was a drift toward increasing economic regulation. This took many forms: direct government regulation of business, union involvement in wage-setting and work organization, restrictions on trade, minimum and maximum prices and wages, public enterprise, and many others. In general, they promoted greater equality of opportunity and outcomes across large sections of society, and they were often promoted with exactly this in mind.

Beginning in the early 1980s this trend was reversed in different ways and to different degrees in different countries. In some unions played a much reduced role; in others the economy was opened to greater international competition. Publicly owned firms were privatized, and market regulations were scaled back. Within firms as well, market-based competition began to replace more administrative systems of control, as we saw in Chap. 8. If it is true that competition is intrinsically inequality-producing, we would expect these political and economic changes to result in the sort of trends that have actually materialized.

It is probably too much to say that competition *always* has this effect, but economists have closely studied the effects of specific forms of deregulation, and

there is general agreement that they explain at least some portion of the overall rise in inequality. When it comes to how much, the agreement breaks down.

4. Winner-take-all markets. Most competition occurs at the margin, as competitors try to acquire a little more or a little better. Companies seek to win market share from one another, and they measure success in percentage points. Workers compete for the best jobs, and success means getting a somewhat higher position, more money or other benefits than those who lose out. As we are used to thinking of it, the world is not divided between total winners and total losers, but between those who have gained a bit more and those who must settle for a bit less.

Some observers think that new forms of technology are changing this pattern. In many fields digital reproduction combined with faster and more thorough communication are creating conditions in which small advantages in productivity, quality or reputation can translate into enormous differences in outcomes. For instance, as financial deals between companies become larger and more complex, the impact of a slightly more skilled lawyer or analyst can amount to hundreds of millions of dollars, and this naturally shows up in the form of a very large income differential between the superstars and those in the middle of the pack.

Network externalities, which we surveyed in Chap. 15, also contribute to this effect. It is common for a particular software or internet-based service firm, for example, to capture nearly all of the available market by being the first to enter it or by having some small advantage over its competitors. This makes the managers and creative people who are little more effective in the speed and quality of their projects vastly more valuable to the companies that employ them.

Finally, digital reproduction and enhanced communication tends to concentrate market share for a wide variety of professionals. Top-level physicians now offer their diagnoses in difficult cases that arise halfway around the world; the designer of a new print font can have millions of users within the space of a few months. This is fine for those whose work is replicated or extended to faraway locations, but it diminishes the contribution of local, less exalted producers.

These are the main theories that purport to explain why this is an age of increasing inequality. Research on their relative importance is inconclusive, and it may well be that the actual processes at work are simply too complex to be described at this level of generality. Nevertheless, they provide the main basis for developing policies to limit or channel inequality in market economies.

Before sketching the most prominent policy alternatives, it should be mentioned that not everyone agrees that policies are needed in the first place. There is political debate over what criteria should be used to determine whether inequality has become too large, as we will see in the [Appendix](#) to this chapter. Also, many people benefit from this inequality, and it is in the nature of our political systems that those with the greatest economic resources also play a disproportionate role in setting the political agenda. So consider the list that follows as a guide to what can be done about inequality *if* there is public support for moving in this direction.

1. **Enhancing human capital.** If skill-biased technical change is the main reason for the upward trend in inequality, then spreading the necessary skills more evenly through the population is the logical antidote. Supporters of this approach would spend more money to improve public schools at all levels, paying particular attention to those who would otherwise fall behind or drop out. Early childhood programs are also favored, since there is evidence that learning skills are strongly affected by experiences children have before they enter school. Proponents of the human capital approach are confident that, if many more young people are given the skills to do intellectually demanding work, more jobs requiring those abilities will materialize; others are more skeptical.
2. **Redirecting globalization.** If intensified global competition is largely responsible for increased inequality, then the speed or direction of globalization may need to be altered. Here there is a wide variety of proposals: reduce the volume or velocity of capital investment flowing between countries, install financial mechanisms that would promote more stable and balanced trade, include labor standards in trade agreements. These ideas flare up each time a new trade agreement is drafted.
3. **Reregulation.** If deregulation is the culprit, then some of the mechanisms of regulation that have fallen into disuse could be resurrected, or new forms developed. Here too the specific ideas are varied: instituting or raising statutory minimum wages, changing the legal environment to make it easier to form labor unions, giving unions or other worker organizations more say in the running of enterprises, promoting other forms of stakeholder influence in business decision-making (as discussed in Chaps. 8 and 17), and using tax or subsidy incentives to encourage more egalitarian practices (or discourage very large privileges for those at the top). In practice, debate around these measures tends to center on whether they would impose an economic price in the form of less dynamism and slower growth.
4. **Redistributive taxation.** If the ultimate sources of the surge toward inequality lie in the nature of modern technology, and if the development of this technology is beyond political control (which may not be the case given the large role of publicly funded research), then great differences in the rewards dished out by the market will simply be a fact of life. The only way to offset this trend would then be to redistribute some portion of these unequal incomes through the tax system, by increasing the tax rate on the highest incomes or financial assets and transferring more of the revenues to the bottom. All industrialized countries already do this to some extent; the question is whether they should do so even more. Economists disagree on whether we should expect positive or negative growth effects from greater redistribution; in any case, the impact in isolation from other policies is not likely to be very large. The main question is whether such leveling is justifiable, an issue we will explore in more detail in the [Appendix](#).

18.7 Discrimination

In every society there are groups which are at risk of being treated unfairly in the economy. In some places ethnic and racial minorities are in this position; in others the distinguishing factor may be religious or caste-based or the result of immigration or resettlement. Skin color often plays a role, and in most of the western hemisphere and parts of Asia indigenous people are singled out for inferior treatment. Moreover, in every country women struggle for economic equality vis-a-vis men. In speaking of **discrimination** in this section, we will refer to the unequal treatment of people who are equal in all relevant economic respects, and whose lesser outcomes are due to the social group they are part of. If two workers are of equal ability, but one is paid more than the other because of her race or national background, this would constitute discrimination—but so would paying them the same if the one from the disfavored background was significantly *more* able or productive, and if this group affiliation were the reason. Thus the definition of discrimination brings into play two elements, comparisons of treatment and the role of group membership. It is important to bear in mind that there is no agreement on which groups should be monitored for potential discrimination, so the definition will apply differently based on how it is interpreted. For instance, is it discriminatory to favor job applicants who are tall compared to those who are short, even if they are equally qualified for the work? There is some evidence that favoritism of this sort exists, but whether it is viewed as discrimination, and therefore a social problem, depends on whether we care about equal treatment among those of different heights.

To organize our thinking about discrimination, it will be helpful to do a simple algebraic exercise. Suppose there are two groups in society, the Uppers and the Lower, or U and L for short. Suppose also that there are two types of jobs, Good and Bad or G and B. If the wage rate for good jobs is g and the wage rate for bad jobs is b , and if the percentage of Uppers in Good jobs is u and of Lower in Good jobs is l , then the average wage gap can be written as:

$$[ug + (1 - u)b] - [lg + (1 - l)b] = \text{Average wage gap between U and L} \quad (18.1)$$

The first bracketed term on the left-hand side represents the average wage for the U group. A fraction u earns the Good wage g , while the rest $(1-u)$ earn only the Bad b . The second bracketed term represents the average wage for the L group, in which l earn g and $(1-l)$ earn b . Multiplying out the left-hand side and then simplifying gives:

$$\begin{aligned} ug + b - ub - lg - b + lb &= ug - ub - lg + lb \\ &= \text{Average wage gap between U and L} \end{aligned} \quad (18.2)$$

Now factor:

$$(u - 1)(g - b) = \text{Average wage gap between U and L} \quad (18.3)$$

(Multiply out Eq. 18.3 to make sure it's right.) What this final version of the equation tells us is that the average wage gap between the two groups—the degree of discrimination if the groups are equal in economically relevant respects—is the product of two factors, the difference in the percentages of each group who hold the Good jobs ($u-1$) and the difference in pay between Good and Bad jobs ($g-b$). To keep track of this, let's call the first factor the **selection effect** and the second the **reward effect**. Intergroup comparisons reflect both of them taken together.

What theories have economists developed to explain the sort of outcomes we saw in Figs. 18.4 and 18.5 above? Although there are many variations and wrinkles, we will summarize them as falling into three camps: taste for discrimination, statistical discrimination and differential bargaining power. As is often the case in the realm of economic theory, the arguments are not mutually exclusive; all may be true in some instances or even in most.

1. Taste for discrimination. This approach was pioneered by Gary Becker, who received a Nobel Prize for this and similar work. (We encountered him earlier when we explored the concept of human capital in Chap. 16.) He envisions a small business owner who hires a group of workers and supervises their work. This owner is not the single-minded profit-maximizer of traditional economic theory, however; he also derives personal satisfaction from enforcing discriminatory attitudes, such as hiring workers from a favored group even when they are less qualified or paying them more even when their performance is no better. In other words, in addition to having an appetite for profit, this business owner has a “taste for discrimination”.

At its core, this is a rather minimalist analysis: it says that the owner discriminates because he wants to. What makes the theory interesting, however, are its predictions. Suppose this owner faces a perfectly competitive market, where the smallest difference in price can mean failure or survival. (This is what is signified by the perfectly elastic demand curve faced by the competitor in such a market.) By exercising his discriminatory taste, the owner runs the risk of being driven out of business. Hiring a less qualified worker, for example, will increase the cost of production. This would also be the result of paying some workers below and others above the value of their productivity: the first would be recruited away to other companies, while the second would earn more than they contribute. Ultimately, our discriminatory owner would face a choice: either accept a lower rate of profit, shut down the business (and lend out his money to others who earn the average rate of profit), or end his discriminatory ways. Over time, it is reasonable to expect that fewer and fewer would choose option 1, which means that discrimination would gradually disappear from the market.

There is some support for this prediction in real-world economies. Women, as we have seen, have seen their average earnings come closer to parity with men in the United States (and other countries); some countries with extremely rigid forms of job discrimination, such as South Africa under apartheid, have found them expensive to maintain in a globally competitive world. Of course, there are other reasons discrimination has diminished in these instances, above all the organized pressure brought by those being discriminated against.

Becker's theory becomes more open-ended when additional complications are introduced. For instance, markets are often less than perfectly competitive, and this creates space for discriminatory employers to continue to exercise their prejudices. Even if profits are somewhat lower they may remain above the opportunity cost of capital. Also, coworkers or customers may be prejudiced rather than, or in addition to, the owners. Bigoted coworkers, by withholding cooperation, can reduce the productivity of discriminated workers, "justifying" (in an economic but not ethical sense) their lower pay. Bigoted customers can have the same effect if they are less likely to buy from companies that have salespeople or other employees from disfavored groups. There is no reason to expect that these forms of discrimination will be competed away over time.

The prevalence of prejudiced attitudes over a wide portion of society can therefore lead to self-perpetuating discrimination in both the selection and reward effects described above. Groups suffering from discrimination will have less access to the better jobs, and the higher-level jobs that are open to them will tend to pay less than they otherwise would. There is substantial evidence that both effects operate in economies like the US.

2. Statistical discrimination. This theory was first proposed by another Nobel laureate, Kenneth Arrow. As formulated, it pertains mainly to the selection effect, assuming that differences in rewards are determined by other factors. Arrow begins with the observation that discriminatory individuals typically think on the basis of stereotypes: they attribute to all members of particular groups the characteristics that only some of them actually have. For instance, an owner of a trucking business might be less willing to hire female applicants, thinking that women, as a group, are worse drivers than men. Perhaps most of the stereotypes that fuel prejudice are wrong. In this case, it may well be the case that women drivers are no worse than men, or are even better on average. If so, over time we are back in Gary Becker's territory. If the assumption about women drivers is wrong, this will eventually show up in the form of lower profits for companies that discriminate against them and higher profits for those who don't. Market competition will weed out those who don't learn from experience. False stereotypes will be dispelled.

But Arrow recognized a second possibility: what if, in this case, the prejudice is true *on average*? That is, suppose (to continue our example) there is a great range in the quality of drivers among both men and women, but that somewhat more women are bad drivers than men? If the employer had full information about the aptitude of each applicant, it would be no problem to select only those with the highest skill, whether men or women. A somewhat higher proportion of those selected would be

men, but by our definition there would be no discrimination, only rational employment practices.

In the usual case, however, the employer does *not* have this sort of information. Perhaps, for instance, it would be too expensive to give each applicant the sort of thorough driving test that could differentiate the skillful from the klutzy. In that case, it is entirely rational for the employer to discriminate on behalf of men: by doing this they would recruit a higher quality workforce, on average, than they would if they did not take gender into account. The result would be a male bias in the workforce out of all proportion to the true difference in average ability. This would be unfair to female applicants, but there would not be any economic cost to the employer—quite the contrary.

The statistical theory proposes a paradox. In the case of false stereotypes, there is no basis for discrimination and it can be eliminated without economic cost. But what if the stereotype is based on a true difference in the relevant group averages? What should be the tradeoff between fairness and efficient matching between jobs and workers?

3. Differential bargaining power. This theory is perhaps the oldest of all three; it is too ancient and widespread to ascribe to any particular thinker. Nevertheless, it has been given greater clarity with the emergence of game theory as a central analytical tool in economics.

Recall that in the model of bargaining power developed in Chap. 14, the critical variable is each party's outside option, what they would get if there is no agreement and they must go their separate ways. The one who needs the agreement most is at a comparative disadvantage. In simple terms, this model explains why members of groups subject to discrimination have worse outside options on average and must accept inferior bargaining outcomes. This is, as should be clear, primarily a theory of reward effects—why the jobs typically inhabited by members of some groups pay less than those inhabited by others.

One striking example of this form of discrimination is the situation faced by undocumented workers, who are subject to arrest and deportation if they are identified. Their outside option, if they fail to accept an employer's wage offer, may be exactly that, arrest and deportation. This means they have no bargaining power at all and must accept virtually any offer they get.

The more common state of affairs, however, is that the vulnerability of those subject to discrimination results primarily from the fact that this discrimination is widespread. If other employers pay you less, then this means your outside option is worse for the one employer you are bargaining with. This can be a self-reproducing situation, with each inferior outcome making the others "rational". It doesn't embody the optimistic prediction of Becker's main model of the discriminating employer because wages are assumed to be set at least in part by bargaining and not perfect competition.

One interesting implication of the bargaining model is that it gives a potentially important role to wealth inequality as a determinant of income inequality. Normally, if a worker refuses to come to agreement with an employer, this will entail a spell of unemployment while he or she looks for another job. (You may want to

reread the section on searching and matching in labor markets from Chap. 16.) The cushion that makes such a spell bearable is the worker's savings. In fact, savings serve two functions: they make it possible to maintain approximately the same level of consumption during interruptions in income, and they give the worker more time to find a better job, rather than having to settle for the first one that comes along. In both ways they diminish the cost of failing to agree with any particular employer. It is often the case that, behind large differences in average group income, lie even larger differences in average group wealth. This is true, for example, for comparisons between Blacks and Whites in the United States.

Policy remedies for discrimination reflect the different viewpoints we have described, as well as the different interests that result from being on one side of the inequality or another. Here we will briefly look at anti-discrimination laws, affirmative action laws and policies to reduce reward effects in general.

1. Anti-discrimination laws. In many cases it is possible to see discrimination as it occurs. Job announcements may be written in a discriminatory way, pay policies may explicitly favor members of one social group, and prospective workers may be directly told that their applications have been rejected for discriminatory reasons. Such practices can be outlawed for specified aspects of the economy (hiring, pay, promotion, lending, etc.) and specified groups who are to be free of discrimination. These types of laws are common in the industrialized countries and are becoming more common in the developing world. They are particularly effective at addressing discriminatory tastes by employers, retailers, lenders and others. Usually they are written in such a way as to permit at least some forms of statistical discrimination, but exactly how much to permit or prohibit is typically a contentious question.
2. Affirmative action laws. These regulations use statistical indicators as signs of potential discrimination. In some cases this is necessary because it is not possible to see prejudice in action. For instance, employers may not state directly that their reason for hiring only workers of one ethnic group for a particular position is due to discrimination, but if the numerical imbalance is so great that it could not be the result of chance, this could be taken as sufficient evidence of discriminatory intent. In this way, affirmative action laws often have the same purpose as anti-discrimination laws, but with a lower barrier of proof.

A second use of these laws is to push employers (and others in a position of power) toward more equal selection choices, irrespective of motive. For instance, even if there is evidence of "rational" statistical discrimination, it may be felt that the social cost of unfairness outweighs the gains from discriminating. If this is the case, affirmative action is the most direct form of remedy. This could also be a response to discrimination due to the tastes of coworkers or customers, where employers would transmit this discrimination to hiring, pay and promotion even in the absence of any prejudice on their own part. Finally, affirmative action can play a significant role in equalizing bargaining power by providing better outside options for members of groups identified and helped by numerical indicators. The debate over whether such indicators should be seen as "hard" (mandatory) or "soft" (indicative, considered in conjunction with other

factors) is about the relative weight of the costs and benefits of affirmative action in light of other social goals.

3. Reduction in reward effects. Much of the debate over reducing discrimination tends to center on selection effects: how can we fairly allocate the better and worse jobs generated by our economies? It is important to remember, however, that group differences depend on both the selection and reward effects, as modeled earlier in this chapter. Even if there is no change in discriminatory selection, group inequalities can be reduced by making the overall distribution of economic rewards more equal in society. To put it in the opposite way, even though many countries are moving strenuously in the direction of equalizing selection through anti-discrimination and affirmative action laws, they may be failing in their larger purpose of reducing discrimination due to the underlying trend toward greater income inequality overall. This means that the policies considered above pertaining to general inequality should also be considered as antidotes to discrimination.

The Main Points

1. If the world were a single country, the gini coefficient of its income distribution would be about 65, where zero is complete equality and one is complete inequality. This is approximately the upper limit for inequality measures within individual countries. More developed countries generally have lower gini coefficients, but English-speaking countries, other things being equal, tend to have higher ones. Latin America tends to have greater inequality. The United States has seen rising inequality over recent decades.
2. There is a broad tendency for countries that have more equal income distributions to also have greater mobility between income classes across generations.
3. Women in the United States have made slow progress toward wage equality with men in the United States since 1979; movement toward equality for Blacks and Hispanics has stalled.
4. There are several ways to measure inequality, including ratios between the income shares going to upper and lower groups. One of the most comprehensive is the gini coefficient. It measures the ratio of the area between the perfect equality (45-degree) line and the Lorenz curve to the total area beneath the perfect equality line, where the Lorenz curve depicts the increase in the share of total income accounted for as the share of the population is increased.
5. A different way to approach distribution is through the functional distribution of income—the shares received in the forms of wages, interest, profits and rents. Income from assets, which include the last three of these, is concentrated in a minority of the population, since asset ownership is concentrated. Human capital, as measured by education, is more evenly distributed than financial capital, and it also receives a return, which economists have been interested to measure.
6. Four general explanations have been given for the trend over recent decades for inequality to rise: skill-biased technical change (the tendency for new

technologies to increase the productivity of more educated workers compared to those with less education), globalization (which puts many categories of workers in competition with one another), deregulation (intensifying market pressure on both workers and employers), and the rise of winner-take-all markets (where slightly more productive workers capture a large proportion of market earnings).

7. Depending on the importance they attach to these different explanations, economists have looked to various policies to moderate or reverse the trend toward inequality: greater investments in education to reduce gaps in human capital, changes in the rules governing international trade to reduce its destabilizing effects, introducing new forms of regulation into labor and other markets that facilitate gain-sharing, and, if none of these are sufficient, redistributive taxation.
8. Discrimination is the unequal treatment of individuals who are, for economic purposes, equal. The difference between the average outcomes to two different social groups, which often provides the measurement for discrimination, can be broken down into a selection effect and a reward effect. The first represents the difference in the access members of these groups have to desirable economic positions, such as better-paying jobs. The second measures the economic gap between more and less desirable positions. In recent US experience, the economic gap between whites and blacks has widened even though economic barriers to blacks have been slowly falling; the remaining gap in opportunities has become bigger in economic terms due to rising inequality in general.
9. Economists have offered three theories to explain how discrimination arises and is affected by other economic forces: the “taste” theory (employers, consumers or coworkers have a psychological preference for discriminating), statistically-based judgments (employers, consumers or coworkers, lacking knowledge of the individuals they interact with, fall back on what they perceive to be the average characteristics of the groups these individuals belong to), and differential bargaining power (according to which discrimination can persist because those in disfavored groups have inferior default options). These are not mutually exclusive.
10. Policies to counter discrimination include anti-discrimination laws (prohibiting discriminatory behavior), affirmative action laws (setting quantitative targets for the allocation of jobs, student admissions or other sought-after opportunities across different social groups), and policies to reduce economic inequalities in general (to diminish the reward effect).

► Terms to Define

Discrimination
Functional distribution of income
General inequality
Gini coefficient
Lorenz Curve

Reward effects

Selection effects

Skill-biased technical change

Questions to Consider

1. Was it wrong for the officers to cut the rope to the Medusa's raft even if, by staying connected, they were completely unable to paddle? If not, would it be wrong if there were *some* difference in the ability to paddle? How much difference would separate right from wrong? Was the unequal access to lifeboats the "original sin" in this story? Would a more equal access to the lifeboats change how you think about cutting the rope?
2. Based on the data in this chapter, do you view global inequality as a serious problem, separate from average living standards? That is, does it matter how large the *differences* are between incomes around the world separate from the *levels* of income themselves? Does your answer to this question have any implications for the types of policies you would support at the national or international level?
3. Based on the data in this chapter, do you think inequality in the US is too high? On what basis? Does the trend toward greater inequality disturb you? Is your opinion affected by the international comparisons offered in Table 18.1?
4. Much of the political turmoil of the last 100 years or so in the industrialized countries has been centered on the conflict between "labor" and "capital". To what extent do the data in this chapter support the view that this is a central conflict of interest in society? What other actual or potential conflicts do they point to?
5. Have you experienced any of the four forces identified as possible causes of rising inequality in your personal experience at work? Do you know of any examples in the experiences of friends or family members? Which seem the most plausible to you based on what you have seen? Why?
6. Give examples of each of the four policy strategies to reduce general inequality. Which do you support or oppose? Why?
7. Are there any groups in society who are not currently protected (at least in law) from discrimination, but who should be? Which ones? Why?
8. Have you seen any of the three theories of discrimination in action in your own experience, or are you aware of them in the experiences of people you know? Explain.
9. Based again on the actual experiences of discrimination you have witnessed or heard about, what types of policies would be effective against them? Which, in your view, would be unjustified? Why?

Appendix: Theories of Distributive Justice

This chapter is about the positive analysis of inequality; this appendix addresses the normative side. In the end, most debates about inequality—whether it is a problem and what to do about it—center not on disagreements over the facts or how they are explained, but on the principles that ought to govern a “fair” distribution. In philosophy, this is the realm of distributive justice. Here our purpose is not to develop complete theories of fair distribution, weighing the arguments pro and con, but simply to present the core principles and the motivations behind them. The point is to be explicit about the criteria we are using when we pass judgment on economic inequality and to avoid fruitless disputes based on unexamined normative assumptions.

Before continuing, we should be clear that the theories we are about to look at all have one element in common, an attempt to apply logical analysis to the problem of inequality. This is not how many people think about it, however. For most of human history, and into the present as well, it has been more common to see the world as made up of “natural” hierarchies: some groups are viewed as above others by virtue of their parentage, gender, physical appearance, cultural aura or social or religious standing. This is the view that has given us kings, castes, the veneration of elders, racial and religious classifications, rigid gender roles, and tight networks of loyalty and patronage. From this angle, an economic distribution would be regarded as just if it corresponds to whichever hierarchies are seen as natural and proper. Even though this is a common perspective—especially if you acknowledge its subconscious force—here we will put aside all appeals to traditional authority and consider only the rationalist arguments associated with philosophy.

Keeping this commitment in mind, we will take up six general approaches to justice: that any outcome that results from neutral rules is just, that rewards should be in proportion to contribution or to effort, that fairness should depend only on initial equality of opportunity, that outcomes should be as equal as possible, and that the primary restriction on inequality should be that an acceptable minimum be guaranteed to all.

1. Neutrality of rules. For many centuries, the Chinese empire was administered by a corps of mandarins, highly educated scholars and officials. These were the highest positions one could aspire to if one was not born to a noble family. To become a mandarin you had to pass a difficult test; only those who had spent many years in study would be able to do this. Admission to the schools was based on achieving an initial level of skill at reading and writing. Of course, for most of Chinese history, only a small portion of the population ever attained literacy, and this meant that most households—mainly peasants and artisans—had no real chance to place a son (this was an all-male institution) in the mandarin corps. Would people of today accept this system as just?

It was extraordinarily unequal; mandarins lived a life that most of the population could only dream of. There wasn't anything approaching equal access to the schools that funneled children into the mandarin class. If you were poor you were

mostly out of luck, although, on occasion, a poor child did manage to ascend through the system. Nevertheless, one argument in favor of this custom is that the rules were the same for everyone, at least everyone who was male. Nothing in the procedures prohibited poor families from trying to get a child admitted to a school, nor was there any penalty to children from poor families if they made it as far as the final exam. The rules were perfectly neutral.

Even today there are some who argue that, as long as the rules that govern who gets ahead are neutral, making no distinctions between anyone, whatever outcomes result should be accepted as fair. Admission to universities and hiring decisions should be “blind”, taking no account of any information about the candidates other than requirements that apply equally to all. What would be unfair, from this point of view, would be any rule that *did* make distinctions, that applied one set of criteria to this person and another set to someone else.

The great advantage of neutrality as a principle of justice is that it minimizes the opportunity for elites to rig the game in their own favor. As soon as you allow procedures that deliberately give an advantage to some over others, you create incentives for insiders to use them to make sure that the advantage is theirs. In the Chinese example, suppose, instead of having a single exam and scoring system for everyone who wanted to become a mandarin, officials were able to choose which applicant would get which exam or could give extra credit to some, but only some, of the exam-takers. In all probability, the children of those who already had the most wealth and power would be given special advantages, and those who came from poor backgrounds would find their way blocked by barriers they might never see.

The disadvantage of neutrality is that the societies we live in are not neutral. As in China during the empire, today we have wealth and poverty, children with every advantage money can buy and others who are up against enormous odds. If the rules are neutral, nothing will intervene to reduce the inequality we have inherited from the past. In the words of Anatole France, “The law, in its majestic equality, forbids the rich as well as the poor to sleep under bridges, to beg in the streets, and to steal bread.” Do we want the laws to make distinctions—to have one set of laws for the rich and another for the poor? Maybe not, but then we might want something else, other programs, policies or institutions that do more than simply determine what is legal or illegal, but intervene in other ways to limit economic inequalities. What these measures should be is the question the other criteria below try to address.

2. Reward according to contribution. A plausible approach would be to say that what individuals get out of a social or economic system should depend on what they put into it. The main virtue is pragmatic: it provides a strong incentive for each person to make as great a contribution as possible. The result will be more to share for everyone, at least if this theory holds true. Indeed, one of the attractive features of the Market Welfare Model is that, if its conditions are met, individuals and organizations (like firms) will earn from the market exactly what each puts in. Thus, each worker would receive the economic value of his or her marginal product, and if that value is correctly measured by prices—as it would be in a Market Welfare Model world—then the worker’s earnings equal the consumer’s benefit from this

work. If the assumptions of the Market Welfare Model are not fulfilled but we continue to accept reward according to contribution as our guiding principle, we should try to fix the market failures or other impediments responsible for supply curves not representing social costs or demand curves not representing social benefits. And failing that, we should use the benchmarks of an idealized perfect market to help us set prices and regulations that will approximate those of a Market Welfare Model world as closely as possible. This, as you should know by now, is the program laid out by conventional welfare economics; see, for instance, the discussion in Chap. 6. The normative underpinning of this project is the belief that rewards in the economy should be governed by contribution.

The ethical basis for this approach is reciprocity, that each should give and get in equal measure. Nevertheless, there are ethical dilemmas that have weakened the appeal of contribution as the main criterion. One is that the contribution one ends up making to society is strongly influenced by pure chance. A worker who works for a particularly innovative or well-run company contributes more than one who works for an average or less effective outfit. According to the principle of reward for contribution, it is appropriate that the first worker should earn more than the second. Nevertheless, it is often a matter of luck which type of job one ends up getting. Perhaps Worker 1, who is employed by the high-powered company and gets a fatter paycheck, has connections that helped her get through the door, or perhaps she was just at the right place at the right time. Is it fair that she should get the greater reward?

Another issue has to do with native human talent. Some simply have greater physical or mental resources to work with. One worker might have to work twice as hard to have the same effect as another; shouldn't this extra dedication be worth something? Are results all that should matter?

3. Reward for effort. Such concerns have led many to adopt a modified version of the contribution principle in which it is the input of the individual (or group), and not the output, that determines rewards. From an ethical standpoint, we are rewarding what people give of themselves for others, even if chance or the uneven distribution of talent does not always translate this effort into tangible benefits. As a practical matter, this approach has the virtue of providing strong incentives to all members of society to work hard, whether or not their skills are in the most demand. Of course, this virtue is also a vice, since rewarding people for simply trying may not give them enough incentive to apply their energies to the things society places its highest value on. In general, market economies do not generate rewards for effort apart from its role in the ultimate product being produced, so some interference in labor or other markets is required if effort per se is to be recognized.

Here is a conundrum you may want to think about in connection with reward for effort: Suppose there are ten workers working on an assembly line. Each operates a different machine, each operation is equally necessary and each is equally difficult. The line moves at a common pace, so workers can neither increase nor decrease the speed of their work. The output of the line is a flow of products that can be sold for a given amount of money. If reward for contribution is the criterion, each worker

should get the same pay. If, on the other hand, the workers are paid in accordance with their effort, the least capable workers will earn the most, since they have to put in more effort to keep up with the pace of the line. Their extra pay will in turn mean lower pay for the most capable, who are, in effect, being punished for having more strength or agility. Is this fair?

4. Equality of opportunity. Perhaps the world is too complex to measure either contribution or effort with any accuracy. There are so many different kinds of effort to compare, and the ways contributions are combined in real-world production systems makes it difficult to tell just who contributed what. Moreover, perhaps both contribution *and* effort should be recognized, although not in any precise combination. In that case we might be drawn to an approach that says, let all start with an equal chance to succeed, and let effort, contribution and luck determine outcomes however they will. Specifically, the criterion of equal opportunity embodies two elements, that there should be a moment in each individual's life (the "starting point") at which equality should reign, and that the rules that govern success should not be biased toward any particular individuals or groups. These are extremely demanding requirements, and it is doubtful that any existing society meets them completely; yet they could serve as goals to be pursued. In practical terms, the first will usually require substantial intervention in market outcomes, since the advantages that children of rich parents would otherwise have need to be offset, but the second is usually thought of as compatible with the way markets should work if they are regulated to be transparent and fair.

There are two large difficulties with equal opportunity as a principle of justice. First, what exactly should be this hypothetical moment of perfect equality—the equal starting line, to use the metaphor of a footrace? Should it be birth? This means that the unequal distribution of luck *before* birth must be counterbalanced, so that those who are congenitally stronger or more clever should be disadvantaged in equal measure. If that doesn't appeal to you, then you perhaps imagine a moment even before birth and before genetic qualities are doled out. But related to this is the problem that, as soon as we are born, we begin to do things or have things done to us that, if not offset, will lead to unequal life chances down the road. The further back we push the moment of equality, the more subsequent inequality we must accept. If opportunities are to be equal at birth, then the advantages that some get in childhood will not count against "equal opportunity". Perhaps you would set a much later age for the "starting point"—say 18. This commits you to much greater intervention to offset all the many pluses and minuses that can accrue by that age, including many that are due to the choices that children make for themselves. At the same time, it can be seen as a bit heartless, since it doesn't allow for second chances. If someone discovers what they truly want at the age of 25 or so, too bad: they missed the moment of equality and they will have to make do with whatever opportunities they are lucky enough to still have. This sort of criticism can be addressed by requiring a multiplicity of "somewhat equal opportunities" that can reappear as one grows older, but then the criterion loses its sharpness: how equal must these second chances be and how many must be offered?

The second large difficulty is that equality of opportunity is compatible with almost any level of general inequality, as we defined it in this chapter. Suppose, for example, that you have a society that works according to this rule: every year a lottery is held with just one winning number. The individual who wins that year gets everything—every last penny of income, all the wealth, the land, everything of value. Everyone else must beg for enough to survive on. The principle of equal opportunity demands only one thing, that the lottery be perfectly fair, so that each person has the same chance to be tycoon-for-a-year, but surely this demand does not go far enough. Can extremely unequal divisions of life's good things be viewed as just simply because the system is fair at the moment just before division?

5. Equality of outcomes. This principle, which we can call egalitarianism for short, does not concern itself with who gets more or how they acquired it; more versus less is itself the problem. The ideal world would be one in which the rewards of our common social and economic life are apportioned perfectly equally. This would presumably be true not at a given moment, such as in equal opportunity, but at all times.

Of course, perfect equality is not really possible. It would take too large a bureaucracy to measure and distribute everything in perfectly equal measure, and we would have the problem of figuring out how much of one sort of good, like security, should be worth another, like income. In practical terms, egalitarianism is an orientation toward more equality of outcomes in preference to less. Thus, to take the case of income distribution, an egalitarian would prefer a more equal distribution (a lower Gini coefficient), all else being equal. It is a preference for equality as a value in itself.

One of the most vexing practical difficulties faced by egalitarians is determining just what it is that should be made more equal. Money income is a logical candidate, but questions arise: Should large families have more income than smaller ones? Should workers with more dangerous or uncomfortable jobs receive extra pay to make up for their hardship? Should people with special needs, for instance due to disabilities, receive the extra income they need to meet them? These are not abstract puzzles in logic; they arise whenever we try to apply egalitarian principles to real human beings.

Another set of problems arises from what egalitarianism is not—not a system that assigns value to contribution or effort. This is an ethical concern, since contribution and effort both make claims on our sense of justice. It is also an issue of great practical importance, since it is very likely that, without the incentives that unequal rewards depend on, people would neither contribute enough work nor be guided to work in the most beneficial ways. To the extent this is true we face what has been called a tradeoff between equity and efficiency, where equity means egalitarianism and efficiency the production of the greatest quantity of economic value. It should be remembered, however, that, while economic incentives are indispensable, they are not the only forces that motivate us, and sometimes they get in the way of more desirable, intrinsic motivators. (See the discussion in Chap. 10.) Thus, the tradeoff between equity and efficiency is real, but not everywhere and always. In any case, it is difficult to claim that egalitarianism should be

the only ethical criterion, and that no weight should be given to contribution or effort, particularly since some degree of inequality will persist in any real-world situation.

6. Guaranteed minimum rewards. In the twentieth century there was renewed interest in approaches to justice that emphasize the needs of those at the bottom of the distribution. Perhaps what we should look for, it was reasoned, is not a principle that governs every issue of distribution for everyone, but which focuses on the crucial needs of the worst-off. There are two main variants at the present time, the brainchildren of John Rawls and Amartya Sen.

(a) Rawls: Maximize the minimum well-being. Rawls, perhaps the most influential philosopher in the field of justice in modern times, takes as his starting point two principles, objectivity and risk aversion. Objectivity means that our evaluation of an economic or social order should not depend on our own place in it; it should be the same no matter which role we come to occupy. His device for achieving this was an imaginary “thought experiment”: suppose we were about enter a society (through birth, for example), but we first had a chance to put an evaluation on it, prior to knowing who we would come to be, including who our parents would be, what our physical inheritance would be, etc. This evaluation would be ideal, according to Rawls, since it would be based on a perfectly objective analysis.

The second principle is that such an evaluation should be governed by risk minimization. What would weigh most heavily in our judgment would be the potential to be someone who is unhappy or oppressed in a society, which we could identify with having a very low income. If this is the case, we would rank societies on the basis of how well-off the worst off person is. If the total income of the economy were a given, Rawls’ theory would dovetail with perfect egalitarianism, since under that rule the worst-off person would be as well-off as possible. Rawls assumes, however, the necessity of economic incentives to motivate production, so that too much equality might well reduce the rewards attainable by the worst-off. The ideal, in his estimation, would be that balance of equality and inequality that maximizes the position of the bottom person in the economic distribution. As a global proposition, this would be difficult to apply, but we might be able to employ it when looking at a particular distributional issue, such as whether the wages for a particular group of workers ought to be raised. Even so, it would be a difficult empirical task to estimate just how much particular distributional adjustments are likely to change economic growth, the extent and direction of innovation, and the like. Toward the end of his life, Rawls retreated from more rigid formulations of his principle.

(b) Sen: Maximize the fulfillment of human capabilities. Amartya Sen follows in the footsteps of Aristotle, who argued that human beings have a common nature and by realizing our potential we can achieve “flourishing”. Aristotle in effect advocated the use of social science to observe a variety of societies to see, empirically, under what conditions their members flourished, so that we could replicate the best of these features in our own ideal. In the more than two millennia that have transpired since Aristotle, however, we have learned much more about how to do

such observation, and we have a much broader base of human possibility to draw from.

Sen believes that we are in a position today to make an informed judgment along Aristotle's lines, provisional (in the spirit of all science) but with real practical implications. We will look at the details of his approach in the next chapter when we focus on poverty, but for now it is enough to mention the broad outlines. Human beings are said to have capabilities, modes of functioning in the physical, psychological and social universe. Exercising strength, solving problems that interest us, engaging with others—these are the sorts of things we all do, but in different ways in different cultures. Sen, who is a Nobel laureate economist as well as a philosopher, suggests that all societies grant “entitlements” to people to enable them to access the resources that make the exercise of capabilities possible. These entitlements may take the form of income, or they may be social obligations to provide particular goods or opportunities, or they may be distributed politically. However we come by them, we need enough of them and in the right combination to fully exercise our capabilities—to flourish.

Although he has not presented his theory in quite this way, it would be a reasonable deduction from his approach to regard the universal attainment of full human functioning as the primary ethical norm. If this goal were achieved, his system of justice would be indifferent regarding the distribution of the “extra” goods not needed for realizing our capabilities. In this way it could be considered a cousin to Rawls, since it would have the primary effect of raising the well-being of the worst off.

This presentation of Sen's theory of capabilities may appear highly abstract, too abstract to be useful in practical situations. This criticism will be addressed in the following chapter, when we see how it has actually been applied, but it should be conceded that this difficulty has never quite been dispelled. There is disagreement among Sen's followers regarding which capabilities are fundamental and how their fulfillment can be measured. There is also a potential risk in any approach which claims to tell us what we “really” need, since economists and philosophers have their own biases and blindspots. On the other (third?) hand, one of the main things we pay philosophers to do is advise us on what we need and how we should regard ourselves.

To summarize, here we have five principles of just distribution. The first three of them try to do too much, and it would probably be a mistake to apply them in every situation. The last (combining Rawls and Sen) at best does too little, since they tell us only about what is fair for those at the bottom, and not for those in the middle or at the top. The fourth (equal opportunity) is both too demanding in its requirements and also not demanding enough. In short, they all have their flaws.

This brief survey should inspire a measure of humility, since whatever yardstick you adopt you will be vulnerable to counterarguments. Also, as in the other truly difficult problems we have explored in this book, you should be encouraged to be flexible, to be willing to use more than one framework when circumstances seem to call for it.

References

- Holzer, H. J., & Hlavec, M. (2012). *A very uneven road: U.S. labor markets in the past 30 years*. Providence: US2010.
- Krueger, A. B. (2012). *The rise and consequences of inequality in the United States*. Delivered to the Center for American Progress, Jan 12. Accessed at <http://www.americanprogress.org/events/2012/01/pdf/krueger.pdf>